

## ECHO CANCELLATION THROUGH AUXILIARY-VECTOR FILTERING

## Background

This embodiment of the invention relates to digital signal processing, and more particularly to echo  
5 cancellation in voice processing systems such as voice mail systems and communication or audio systems.

Echo cancellers are used to reduce the amounts of undesirable echo in telephony voice mail systems. This echo comes from the system's far end as a result of playing voice prompts and is dependent on the distance between the near and the far ends, as well as on the quality of the phone line wires and the connections between them. It is not desired because it can mask a desired response coming from the far end, e.g. dual tone multiple frequency (DTMF) digits. Echo cancellers are usually built as adaptive filters, because they must adapt to the characteristics of specific channels. A digital filter is defined as an array of numbers (also referred to as "coefficients" or "taps"). The array has a length L. A signal is said to be "digitally filtered" when L values of the signal are multiplied by the L filter coefficients.

25 Two problems have been faced with such  
adaptive filters. One is the amount of echo reduction.  
Echo should be reduced to a level which is under a  
specified (desired) level of the DTMF digits, where  
level refers to the power of the signal. Values are  
30 usually required to be under -39 dbm. Another problem  
is the amount of time needed for the adaptation of the  
filter. Echo cancellers' tap values are usually  
initialized either with zeroes or with some initial

values (line probing), but in order to more closely imitate the characteristics of the channel, they need some time to converge, meaning to obtain values that represent these characteristics with a higher precision.

5 There are theoretical scenarios, where user input (DTMF digits) can be missed if the convergence time is too long.

This invention provides the needed echo reduction and reduces the time needed for adaptation of the echo-cancelling filter.

10

#### Brief Description Of The Drawing Figures

Fig. 1 is a schematic block diagram illustrating the echo canceller of one embodiment of the invention in a telephony voice mail system;

15

Fig. 2 is a schematic block diagram illustrating the Auxiliary-Vector utilized in the echo-canceller of one embodiment of the invention; and

20

Fig. 3 is a flow diagram illustrating a method of one embodiment of the invention.

25

#### Detailed Description

Fig. 1 illustrates the echo canceller of one embodiment of the invention in a voice processing system, such as a voice mail system (VMS) 10 for example. A digital communications setup (digital phones, etc.) is considered. Voice mail system 10 comprises a digital signal processor (DSP) 12, phone lines 14 and a voice/signal transducing device in the

30

0035154-04404

form of a telephone set 16. As used herein the one end of the system adjacent the DSP 12 is designated the near end and the opposite end including phone 16 is designated the far end. VMS 10 may also include a voice mail application at the near end. A voice mail application may be, for example, a software program running on a certain type of hardware for example using some of the DSP resources. At the far end there usually is a phone set operated by a human user. There are phone lines connecting the near and far ends. The system 10 may include a first path 24 extending from the voice mail system DSP 12 to telephone set 16. The portion of path 24 at DSP 12 is designated a near end transmitting point 26 (Near end Tx), and the portion of the path 24 at phone 16 is identified as a far end receiving point 28 (Far end Rx). The system 10 may include a second path 30 extending from the telephone set 16 to voice mail system DSP 12. The portion of path 30 at phone 16 is designated a far end transmitting point 32 (Far end Tx), and the portion of path 30 at DSP 12 is designated a near end receiving point 34 (Near end Rx).

The voice mail system at the near end sends voice mail prompts along path 24 to telephone set 16. These and any other signal components from the voice mail system are designated  $x[n]$  in Fig. 1. The signal of interest designated  $s[n]$  is generated by the user at the far end, and one example is a signal representing a button pressed on telephone set 16 when the user hears a voice mail prompt. This signal  $s[n]$  is to be returned along path 30 to the voice mail system at the near end. The undesired echo generated at telephone set 16 is identified by  $x^*[n]$  in Fig. 1. The undesired echo can

003354-04401

be caused in several ways, such as in speaker phone use or accessing voice mail via a cell phone. Noise from the phone lines, making/breaking connections, etc. is represented  $a[n]$  and is added to path 30 at 40.

5

Thus, signal  $x[n]$  represents the voice prompts, played from the near-end, which travel through the telephone network to the far end. A slight amount of echo is present in the received signal  $y[n]$  ( $y[n] = s[n] + x^*[n] + a[n]$ , where  $s[n]$  is the signal of interest, returned from the far end,  $x^*[n]$  is the undesirable echo and  $a[n]$  combines all other noise sources-channel, etc.). In particular, there usually is a slight amount of echoed far end received signal through the far-end transmit (Tx) line. What one would usually like to see in the signal received at the near-end would be a signal which represents the signal from user  $s[n]$ , which may contain user data - voice or commands (usually keys pressed on the phone set are considered commands) - that may be processed by the VMS. An echo canceller 50 is provided and may be operatively associated with the DSP 12 of the voice mail system. The purpose of the echo canceller 50 is to subtract the echo from the received signal, so that only the useful part remains. An illustrative form of echo canceller 50 is software running on DSP 12. The software may be in the form of a program of instructions to process the signals in which undesired signals, e.g., the echo, are cancelled, and the program may be embodied in a program storage device readable by a machine, such as the DSP, and which instructions are executable by the machine

0988154 041401

5

where

15

20

25

30

35

between the real received signal  $y[n]$ , which contains the echo and all other undesired random noise, and the signal  $y^*[n]$ , which is produced at the output of the echo canceller 50 and which is an estimation of all the factors that the real channel superimposes on the received signal. If the signal  $s[n]$  is present at adaptation time, it will be considered undesired and the filter will try to cancel it. So, short adaptation time may be important.

10

In accordance with one embodiment of the invention, the echo canceller 50 employs Auxiliary Vector filtering. It takes the transmitted signal  $x[n]$ , multiplies it by an array of filter coefficients  $w[n]$  and then subtracts the result from the received signal. In the invention,  $w[n]$  is computed by means of an Auxiliary-Vector filtering algorithm. The operation may be summarized as follows:

20

The near end (VMS) starts transmitting.

Signal detection algorithm is started immediately in the VMS. Filter coefficients of the echo canceller are initialized with fixed start-up values.

25

When the echo is detected, the filter starts adapting its coefficient values.

After a certain amount of time, the filter has adapted its coefficients and is ready for the real echo cancellation.

30

The Auxiliary-Vector filtering algorithm may be specified in detail by the following relationships:

35

$$R=Y^T*Y$$

$$V=Y^T*X$$

$$G=(R*V-(V^T*R*V)*V)/\text{norm}(R*V-(V^T*R*V)*V)$$

$$U=(G^T*R*V)/(G^T*R*G)$$

$$w=V-U*G$$

09335154-044401

The vector Y consists of one row and N columns and contains N values of the signal y[n]. In other words N is the number of samples the VMS can process at a given moment. N may also be the echo cancellation  
5 filter length, e.g., the number of filter taps in the array. In one embodiment of the invention, N usually is between 16 and 512.

The first operation in the above algorithm is  
10 to form the matrix R from the vector Y and  $Y^T$ .  $Y^T$  denotes a matrix transposition operation, performed on the vector Y. The matrix R has dimensions N x N. The next step in the algorithm is to form the vector V by multiplying  $Y^T$  on with X which is a single value (or a  
15 vector with one row and one column). X contains one value of the signal x[n]. The resulting vector V has the dimensions N x 1.

The next operation in the algorithm is to form  
20 the Auxiliary Vector:

$$G = (R*V - (V^T*R*V)*V) / \text{norm } (R*V - (V^T*R*V)*V)$$

where the operation norm ( ) denotes a matrix  
25 normalization operation. The resulting Auxiliary Vector G has dimensions N x 1. This is the vector that maximizes the magnitude of the cross correlation between the main-stream processed data (the V branch designated  
50 in Fig. 2) and the auxiliary-vector beam processed data (the U\*G branch designated 52 in Fig. 2).  
30

The next operation in the algorithm is to form the scalar:

090354.04404

$$U = (G^T * R * V) / (G^T * R * G)$$

U is a scalar which minimizes the output variances of the filter coefficients. The next operation in the algorithm is to form the vector:

$$w = V - U * G$$

The vector w consists of one row and N columns and it contains the filter tap values. After the foregoing adaptation process ends, w contains the values which should cancel the undesired noise factors, e.g., the echo.

In particular, as shown in Fig. 3, in stage 70 of the process the input values are obtained and then the foregoing algorithm operations 72, 74, 76, 78 and 80 may be performed. The filter 50 stops the adaptation process after the values of the error signal  $e[n] = y[n] - y^*[n]$  reach some predetermined level as indicated in stage 82. If not, the process is repeated as indicated by line 84 including taking additional samples of  $x[n]$  and  $y[n]$ . This is repeated as many times as necessary including taking multiple samples of  $x[n]$  and  $y[n]$ . Then, the filter 50 operates to remove the unwanted components (noise) from the received signal as indicated in stage 86. In the voice mail case, filter 50 may be adapted on a per-call basis, because the noise may depend on the lines used for the connection, as well as on some other random factors, which, however, are fairly stable on a per-call basis.

The algorithm may reduce the amount of time needed for training the filter 50, therefore reducing



the possibility of missing user input (DTMF digit(s)) in these initial moments. It is believed that in most cases a single run through the process illustrated in Fig. 3 will be sufficient without the need for the  
5 repeat indicated by line 84. Another advantage is that the filter 50 can adapt its coefficients using a smaller number of known (training) samples of data.

While the embodiments of the invention have  
10 been describe in detail, that is for the purpose of illustration, not limitation.

09835154.044401  
FOI4045E860